

We claim:

1. A method of identifying a region of genomic DNA associated with a phenotypic trait of interest comprising:

- (a) screening a mapping population of *Arabidopsis thaliana* plants to determine the linkage of said phenotypic trait with a collection of polymorphisms, wherein said at least one polymorphism is selected from Table A; and
- (b) identifying said region of genomic DNA associated with a phenotypic trait based on linkage of said trait to one or more of said polymorphisms.

2. A method according to claim 1 wherein said mapping population of *Arabidopsis thaliana* plants is screened to determine the linkage of said phenotypic trait with a collection of nucleic acid molecules which are capable of detecting a set of polymorphisms, where the polymorphisms are distributed throughout the *Arabidopsis thaliana* genome at an average density of more than one polymorphism per about 100kb.

3. A method according to claim 1 further comprising calculating the linkage of each of said polymorphisms to said phenotypic trait.

4. A method according to claim 1 further comprising isolating said identified region.

5. A method according to claim 1 wherein said region comprises a gene linked to said trait.

6. A method according to claim 1, wherein screening a mapping population comprises:

(a) identifying an *Arabidopsis* plant of a first ecotype with a phenotype of interest;

(b) crossing said *Arabidopsis* plant with an *Arabidopsis* plant of a second ecotype lacking said phenotype;

(c) propagating and self pollinating seeds from said cross;

(d) selecting progeny of self pollinated seeds with said phenotype;

(e) screening progeny of self pollinated seeds with said phenotype with a said collection of nucleic acid molecules which are capable of detecting a set of polymorphisms.

7. A method according to claim 1 wherein said region of genomic DNA associated with said phenotype is located between about 5 and about 10 cM of one or more of said polymorphisms.

8. A method according to claim 1 wherein said region of genomic DNA associated with said phenotype is located between about 0 and about 5 cM of one or more of said polymorphisms.

9. A method according to claim 1 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 25 polymorphisms selected from Table A.

10. A method according to claim 9 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 50 polymorphisms selected from Table A.

5 11. A method according to claim 10 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 75 polymorphism selected from Table A.

12. A method according to claim 11 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 100 polymorphisms selected from Table A.

10

13. A method according to claim 12 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 200 polymorphisms selected from Table A.

14. A method according to claim 13 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 500 polymorphism selected from Table A.

15

15. A method according to claim 14 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 1000 polymorphisms selected from Table A.

20 16. A method according to claim 15 wherein said collection of nucleic acid molecules is capable of detecting a set of greater than 2000 polymorphisms selected from Table A.

17. A method for identifying transposons in the DNA of an *Arabidopsis thaliana* plant comprising identifying INDELs in said DNA and comparing the sequence of said INDELs to the sequence of one or more known transposons.

18. Computer readable medium having recorded thereon at least 100 of the polymorphisms set forth in Table A.

19. A collection of non-identical nucleic acid molecules capable of detecting polymorphisms in an *Arabidopsis thaliana* mapping population, wherein said collection of non-identical nucleic acid molecules is capable of detecting at least 25 distinct polymorphisms selected from the group consisting of Table A.

20. A collection according to claim 19 wherein non-identical nucleic acid molecules is deposited on a substrate.

21. A collection according to claim 19 wherein each of said non-identical nucleic acid molecules comprises

- (a) at least 12 nucleotide bases and a detectable label, and wherein the sequence of said at least 12 nucleotide bases is at least 90 percent identical to either strand of a segment of *Arabidopsis thaliana* DNA having a sequence which has the same number of consecutive nucleotides as said molecule and which includes or is adjacent to the locus of said polymorphism; and wherein said segment is located

in the BAC which is identified in Table A as having said polymorphic sequence;

or

(b) at least 15 nucleotide bases, and wherein the sequence of said at least 15

nucleotide bases is at least 90 percent identical to either strand of a segment of

5 *Arabidopsis thaliana* DNA having a sequence which has the same number of consecutive nucleotides as said molecule and which includes or is adjacent to the locus of said polymorphism; and wherein said segment is located in the BAC which is identified in Table A as having said polymorphic sequence.

10 22. A collection according to claim 21 wherein said sequence identity is at least 95%.

23. A collection according to claim 21 further comprising for each of said non-identical nucleic acid molecules a pair of isolated nucleic acid molecules useful for PCR amplification of a segment of *Arabidopsis thaliana* DNA comprising at least one

15 polymorphism, wherein each nucleic acid molecule of said pair comprises at least 15 nucleotide bases and wherein the nucleotide sequence of one of said molecules is at least 90 percent identical to one strand of a segment of *Arabidopsis thaliana* DNA having a sequence which has the same number of consecutive nucleotides as said molecule and which includes or is adjacent to the locus of said polymorphism; and wherein said
20 segment is located in the BAC which is identified in Table A as having said polymorphic sequence; and the sequence of the other of said molecules is at least 90 percent identical to the complementary strand of said segment.

add
D2